Reducing Computational Time of Closed-Loop Weather Monitoring:

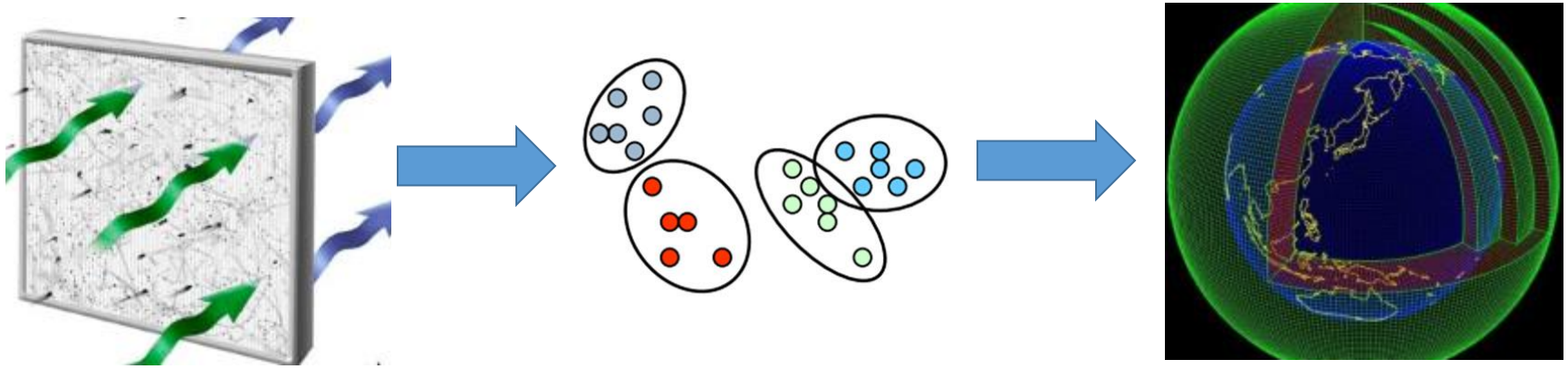
A Complex Event Processing & Machine Learning Based Approach

H. M. C. Chandrathilake, H. T. S. Hewawitharana, R. S. Jayawardana A. D. D. Viduranga,
H.M.N. Dilum Bandara, Suresh Marru[#], and Srinath Perera[*]

Dept. of Computer Science & Engineering, University of Moratuwa
[#]Indiana University, Bloomington, USA
[*]WSO2 Inc., Sri Lanka

Contribution

We attempt to address drawbacks of current weather forecasting approaches by pre-processing weather data using Complex Event Processing and Machine Learning



CEP based filter

ML based clustering module

Complex, weather processing module

Introduction

Weather forecasts…

LARIDAE
ENHANCED WEATHER MONITORING SYSTEM

Very important and need to be **timely** and **accurate**

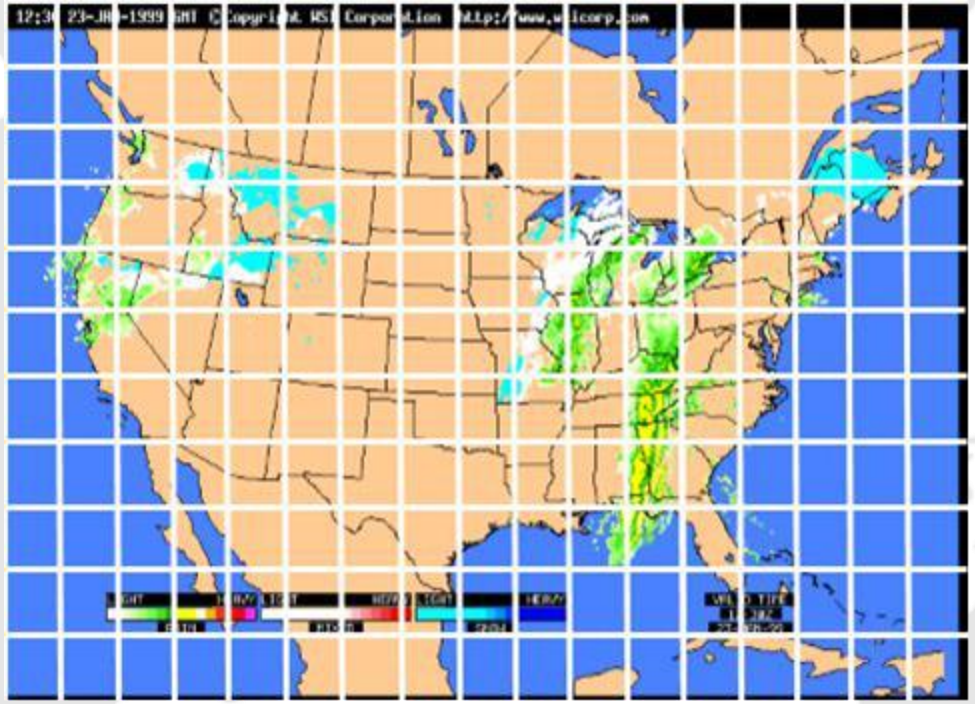
But weather algorithms have to process **too much data**

The Real Problem

LARIDAE
ENHANCED WEATHER MONITORING SYSTEM

Data collection is,

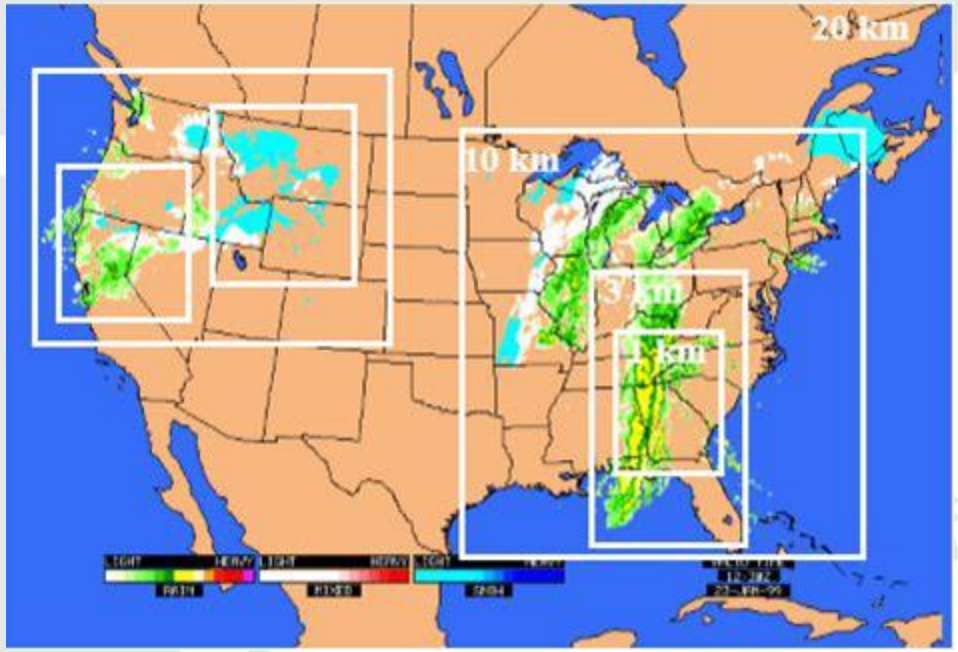
- **Static**
- **Fixed-Cycle**
- **Low Resolution**



The Answer!

LARIDAE
ENHANCED WEATHER MONITORING SYSTEM

- Identifying interesting areas if any
- Increasing the data collection resolution
- Increasing the data collection frequency



Solution Architecture

High Level Solution…

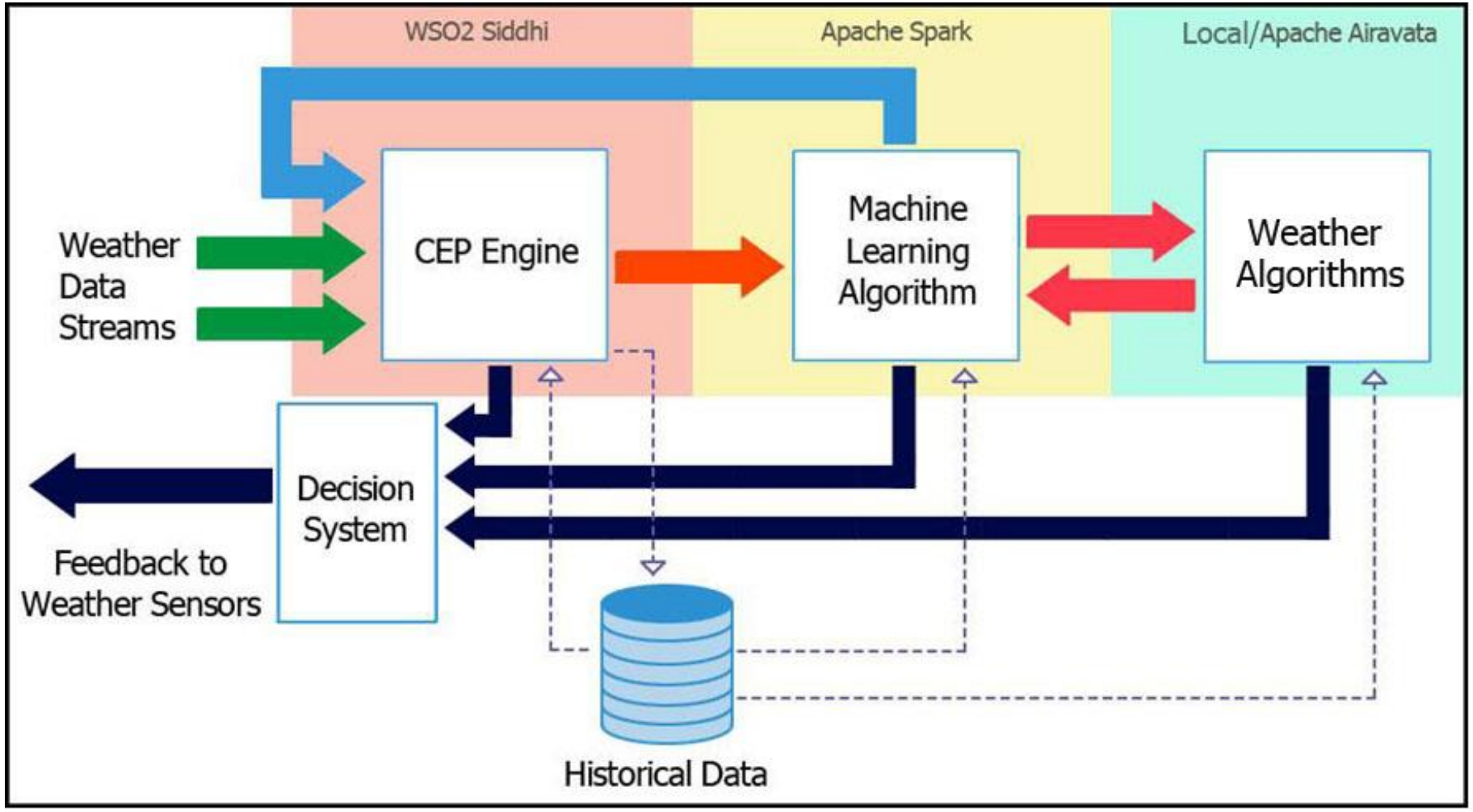
LARIDAE
ENHANCED WEATHER MONITORING SYSTEM

2 Solution Architecture

3

4

5



Selected Case

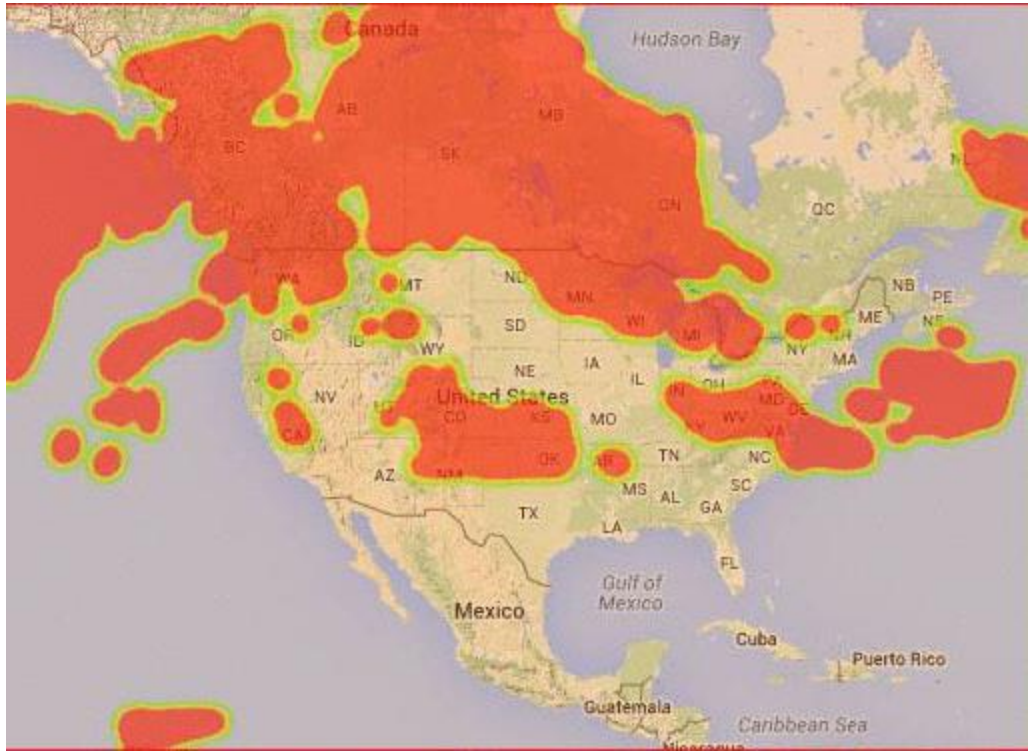
LARIDAE
ENHANCED WEATHER MONITORING SYSTEM

2 Solution Architecture
3
4
5

2006 December, 19-20

Decatur, United States

3 -Hourly Data in GRIB* format

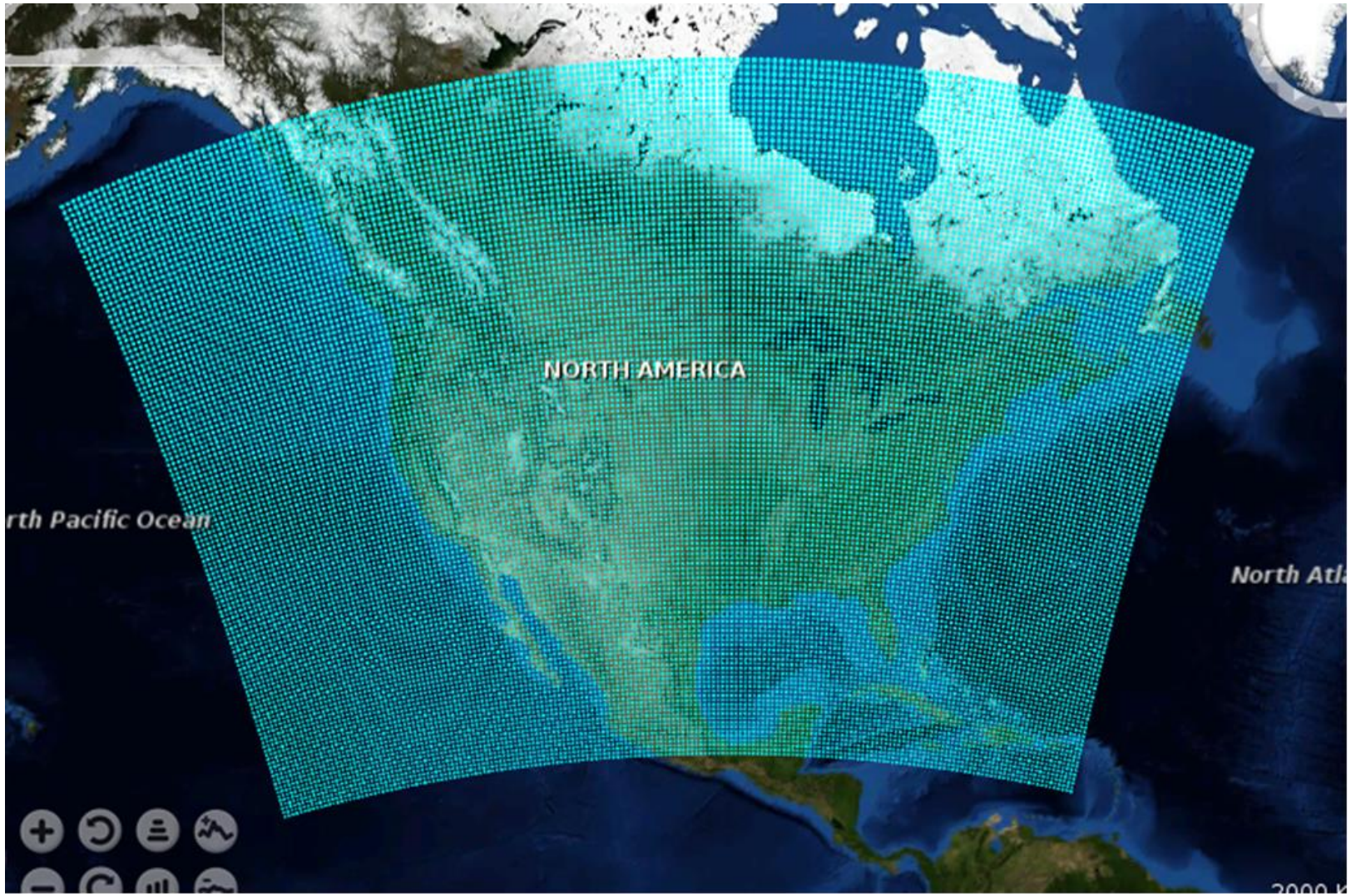


* GRIB is a data format dedicated to handle weather data

Initial Data Coverage

LARIDAE
ENHANCED WEATHER MONITORING SYSTEM

2 Solution Architecture
3
4
5



High Level Solution…

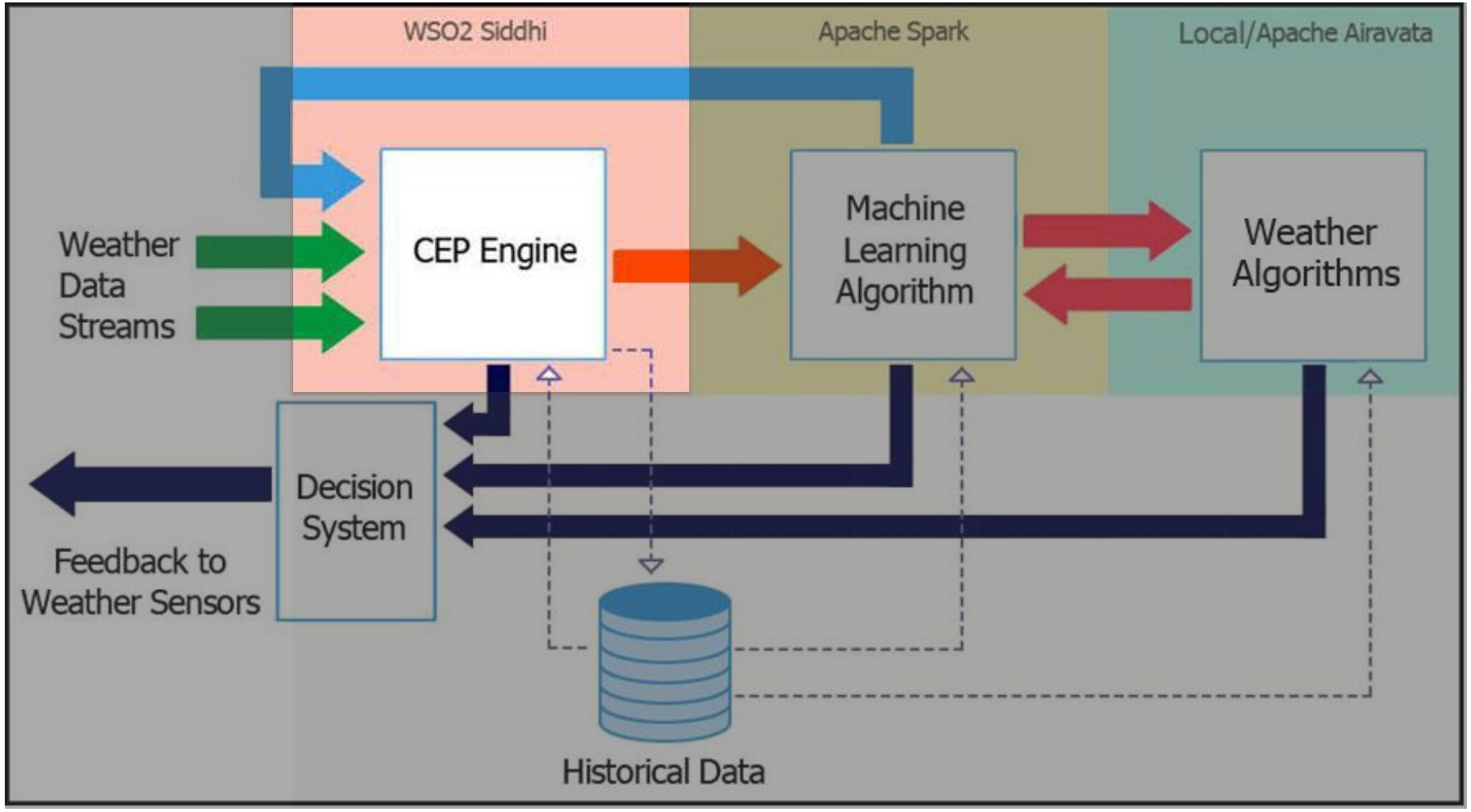
LARIDAE
ENHANCED WEATHER MONITORING SYSTEM

2 Solution Architecture

3

4

5



CEP Filtering Layer

LARIDAE
ENHANCED WEATHER MONITORING SYSTEM

2 Solution Architecture

3

4

5

Complex Event Processing

- Monitoring data streams in real time
- Identifying complex relationships among events (data)
- We used it to monitor weather stream data

Goals

- Anomaly detection
- Filtering

Filtering using weather indices

- Lifted Index
- Storm Relative Helicity
- Convective Inhibition

Lifted Index

LI ≥ 0K : *Stable atmosphere - no thunderstorms possible*
0K > LI ≥ −2K : *Thunderstorms possible*
−2K > LI ≥ −6K : *Thunderstorms likely*
−6K > LI : *Severe thunderstorms likely*

CEP Filtered Data

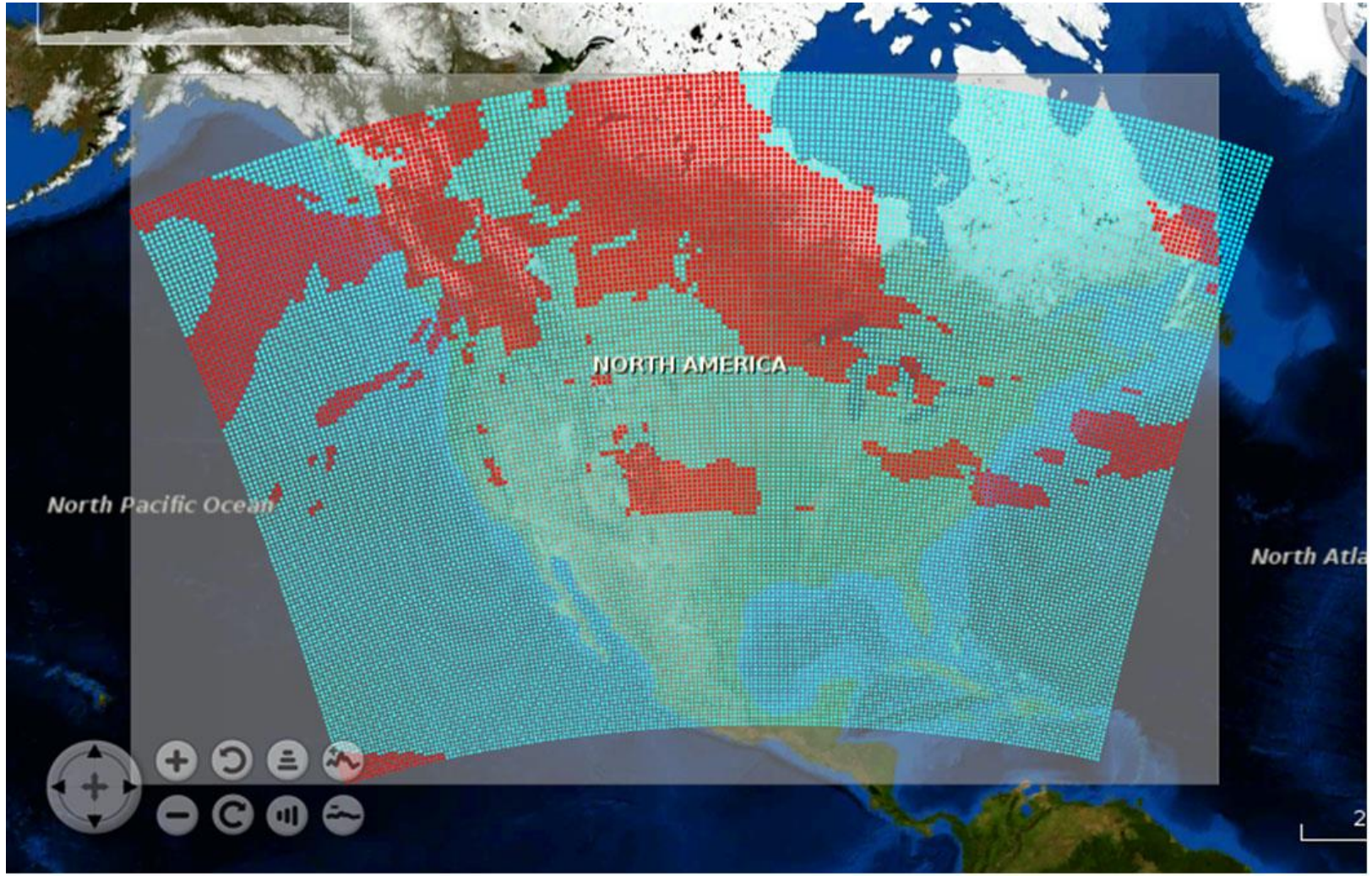
LARIDAE
ENHANCED WEATHER MONITORING SYSTEM

2 Solution Architecture

3

4

5



High level solution…

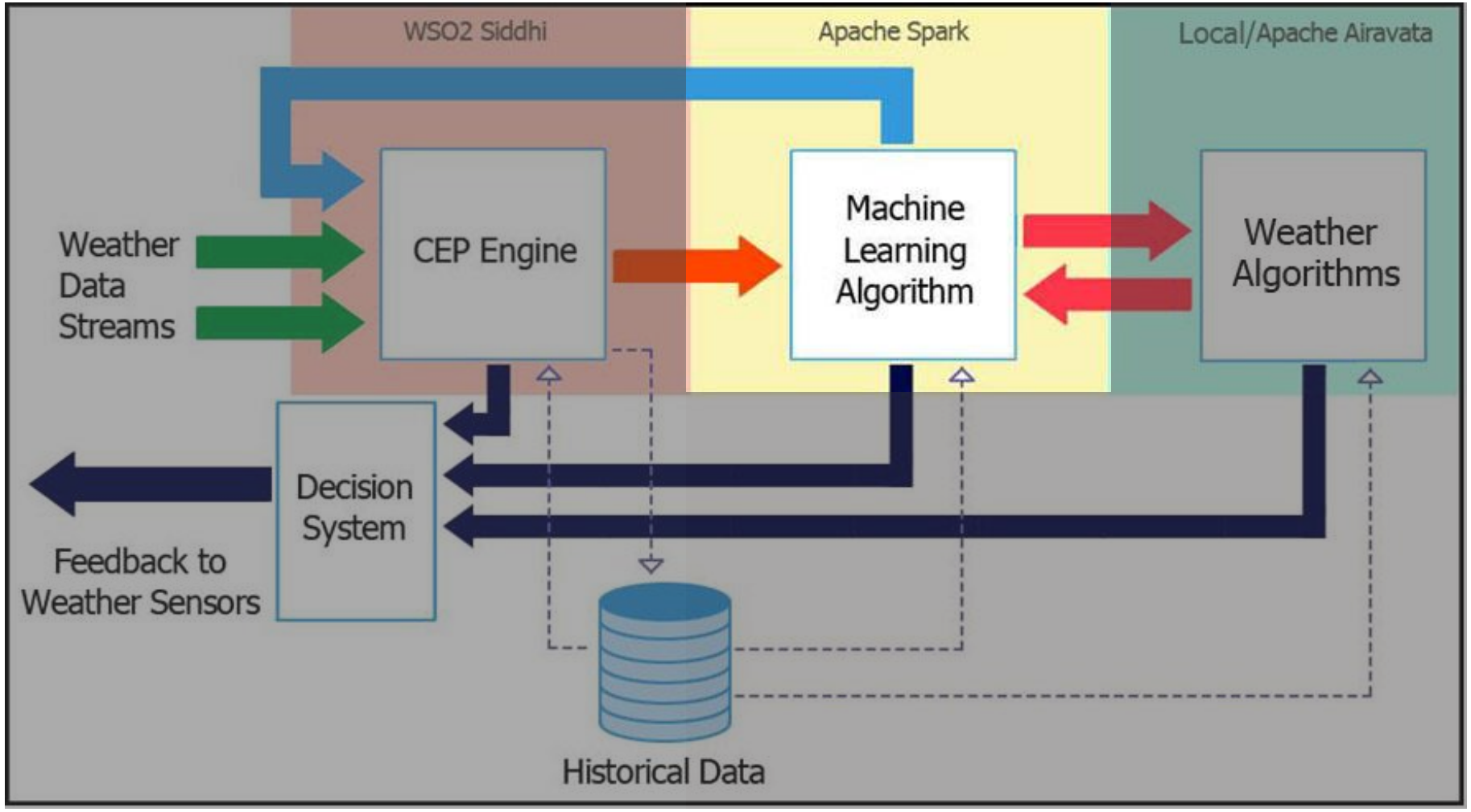
LARIDAE
ENHANCED WEATHER MONITORING SYSTEM

2 Solution Architecture

3

4

5



ML Clustering Layer

LARIDAE
ENHANCED WEATHER MONITORING SYSTEM

2 Solution Architecture

3

4

5

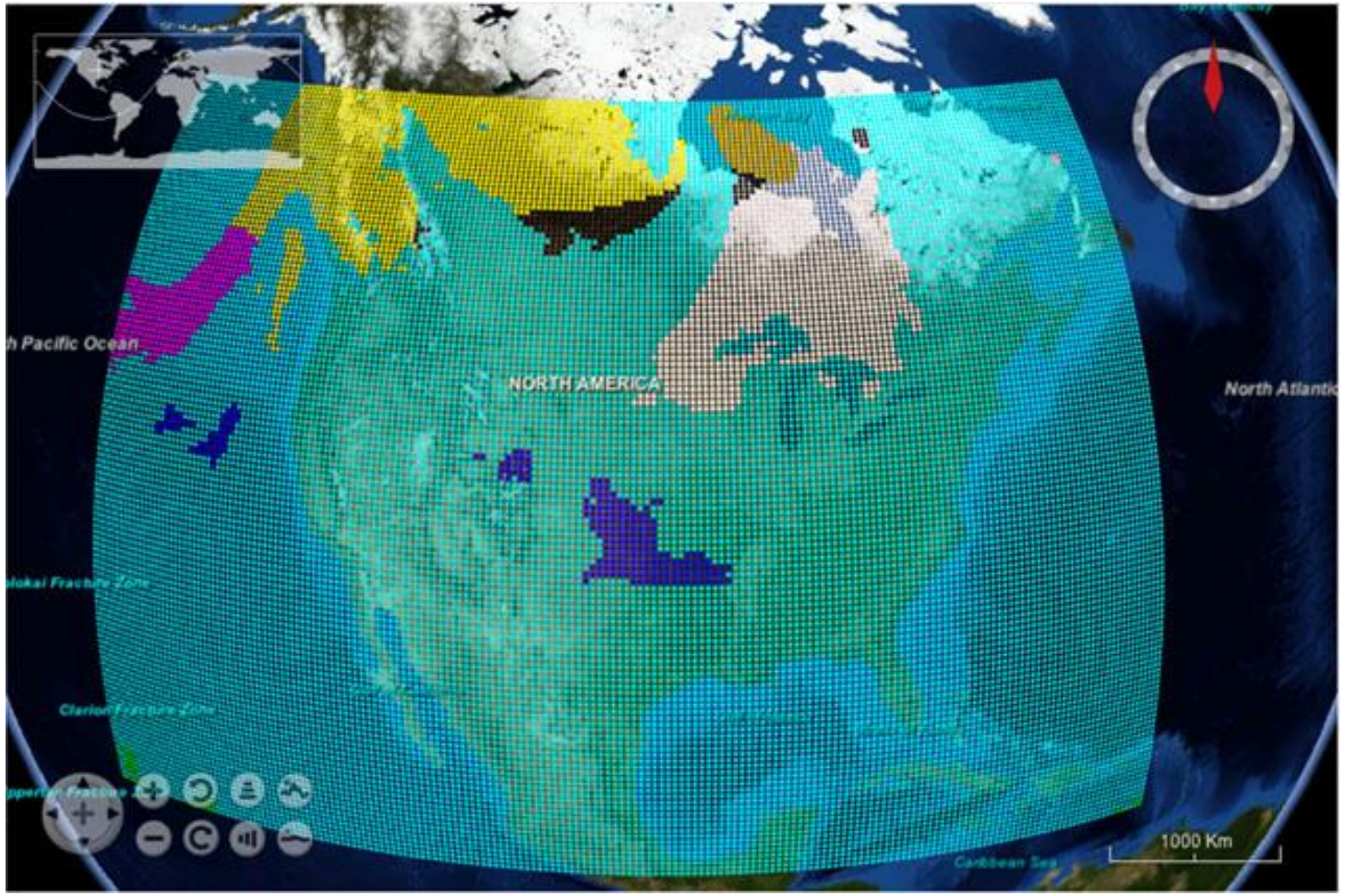
Used to

- Cluster the data points
- Get localized boundaries

2 Algorithms used

- K-means
- Gaussian Mixture Model

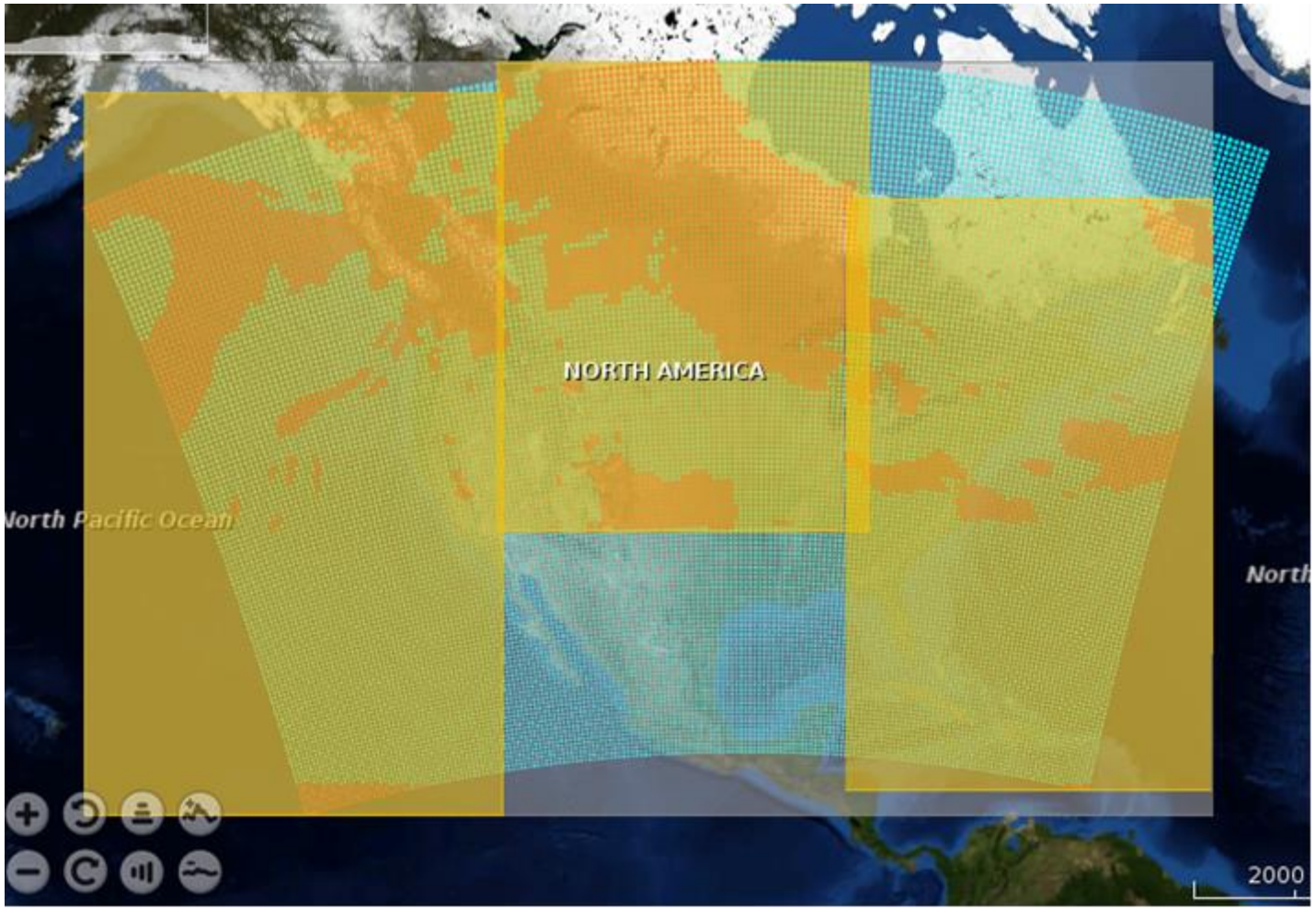
Apache Spark



K-Means Results

LARIDAE
ENHANCED WEATHER MONITORING SYSTEM

2 Solution Architecture
3
4
5



GMM Results

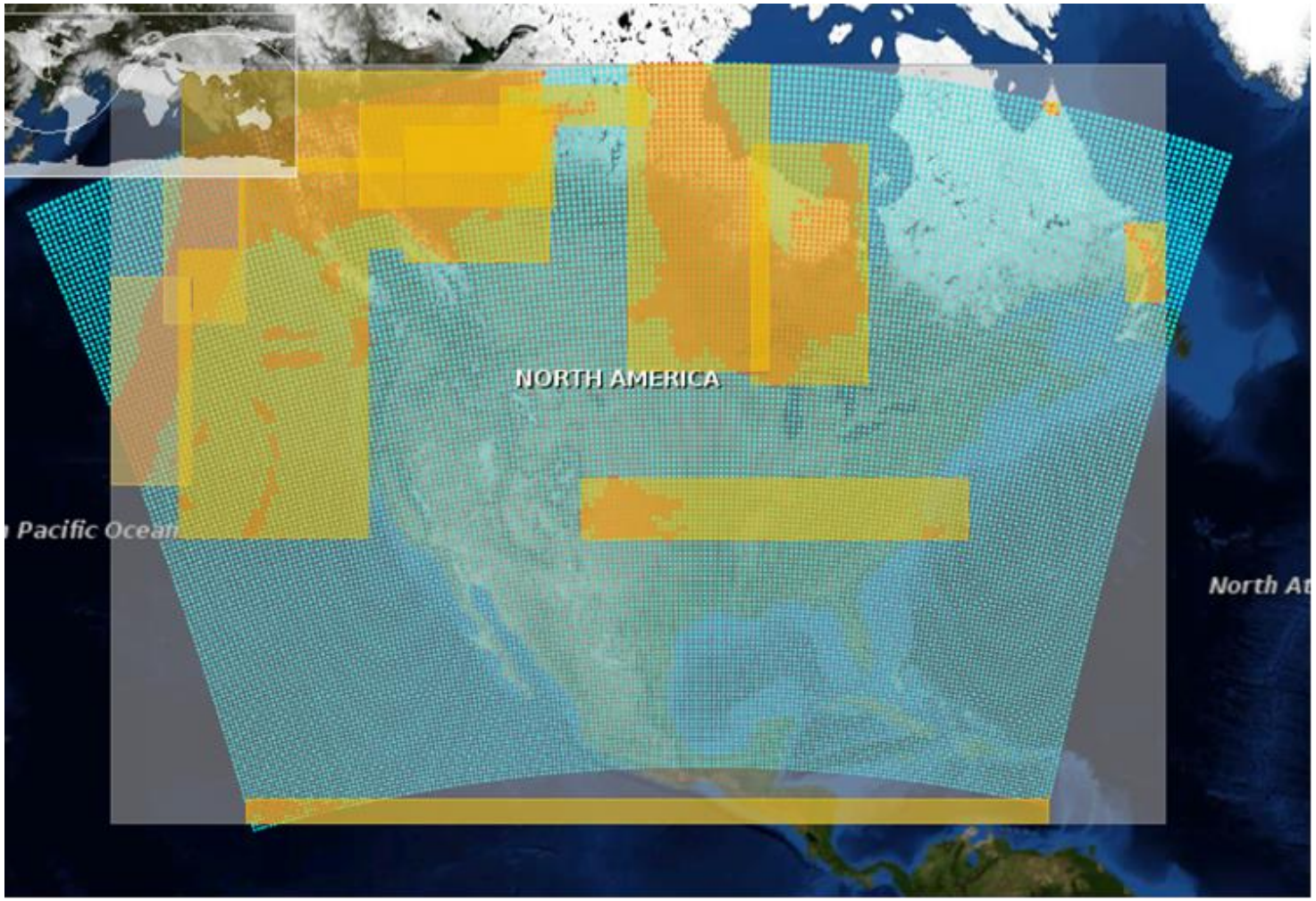
LARIDAE
ENHANCED WEATHER MONITORING SYSTEM

2 Solution Architecture

3

4

5



High level solution…

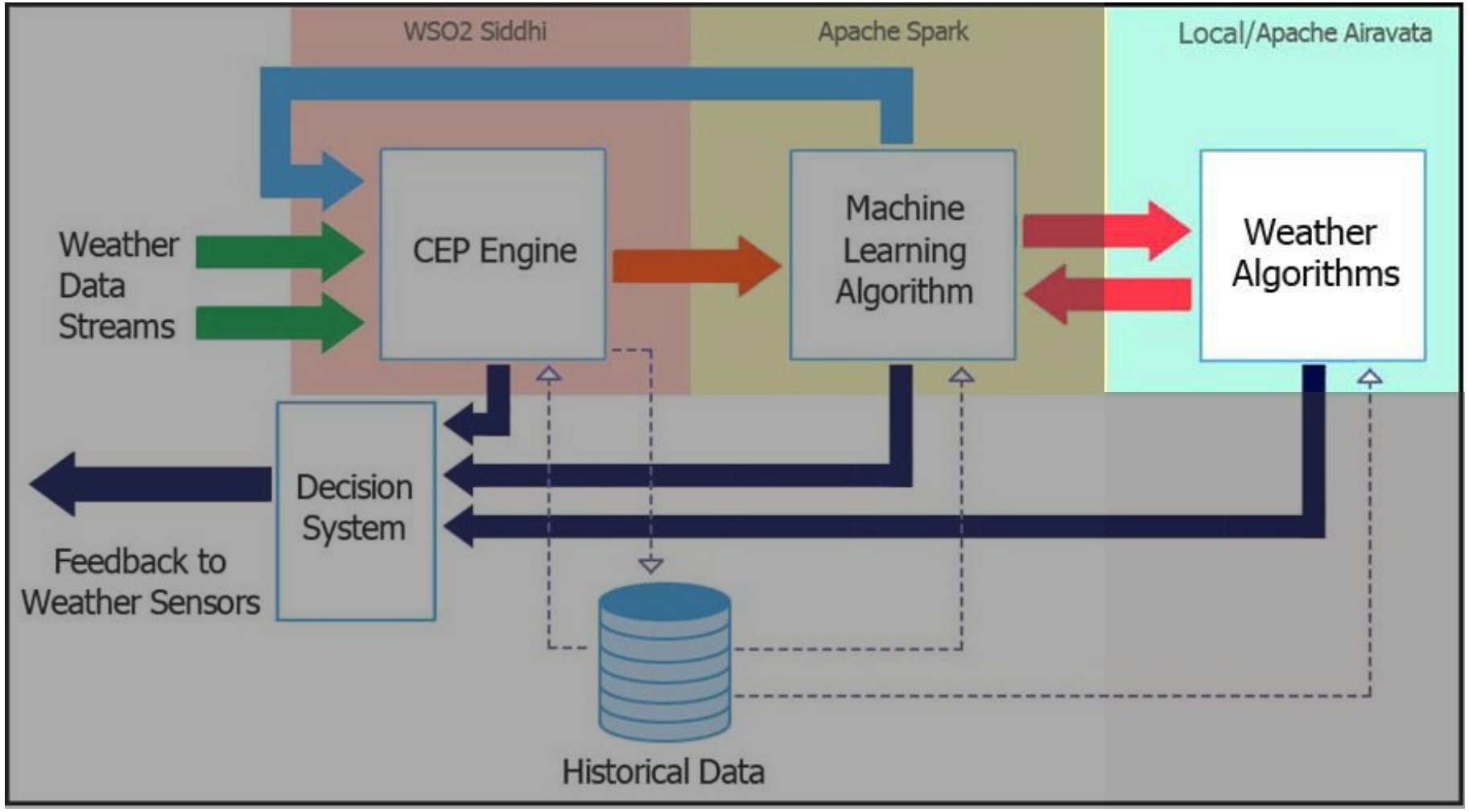
LARIDAE
ENHANCED WEATHER MONITORING SYSTEM

2 Solution Architecture

3

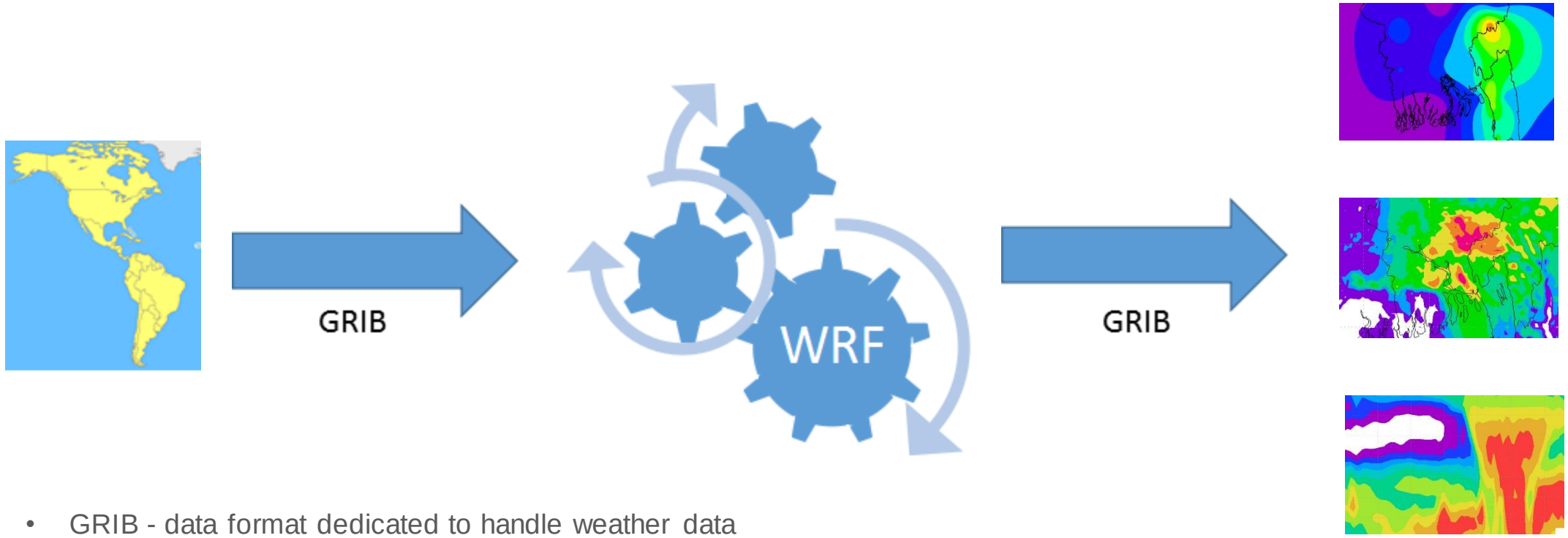
4

5



WRF Model

LARIDAE
ENHANCED WEATHER MONITORING SYSTEM



- GRIB - data format dedicated to handle weather data
- GRIB data provided to WRF model
- Output is received in GRIB/netCDF
- Post processing to get insight

WRF Model

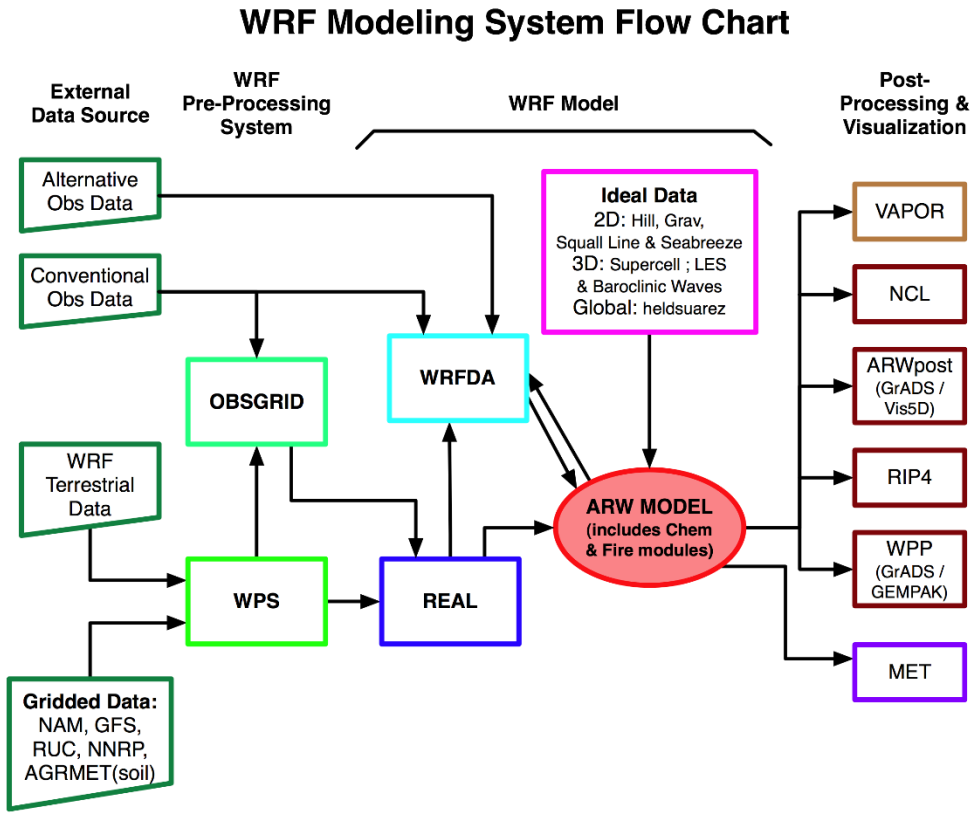
LARIDAE
ENHANCED WEATHER MONITORING SYSTEM

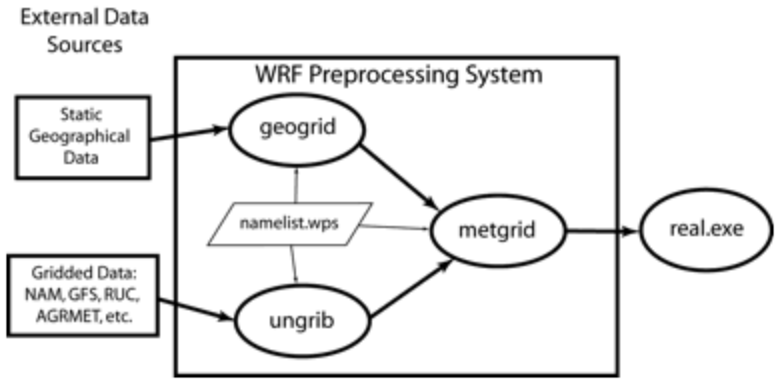
2 Solution Architecture

3

4

5





WRF Model

LARIDAE
ENHANCED WEATHER MONITORING SYSTEM

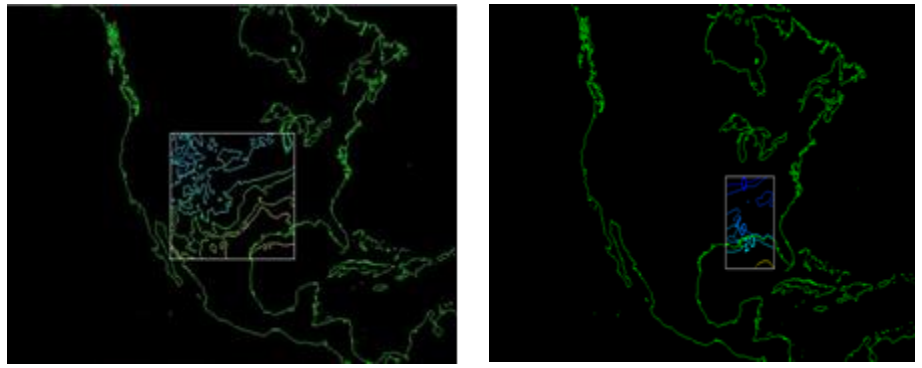
2 Solution Architecture

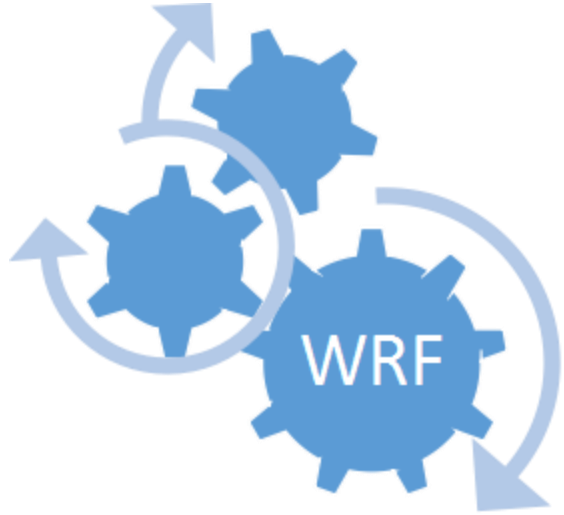
3

4

5

Namelist Configuration





PERFORMANCE EVALUATION

CEP Based Filtering

LARIDAE
ENHANCED WEATHER MONITORING SYSTEM

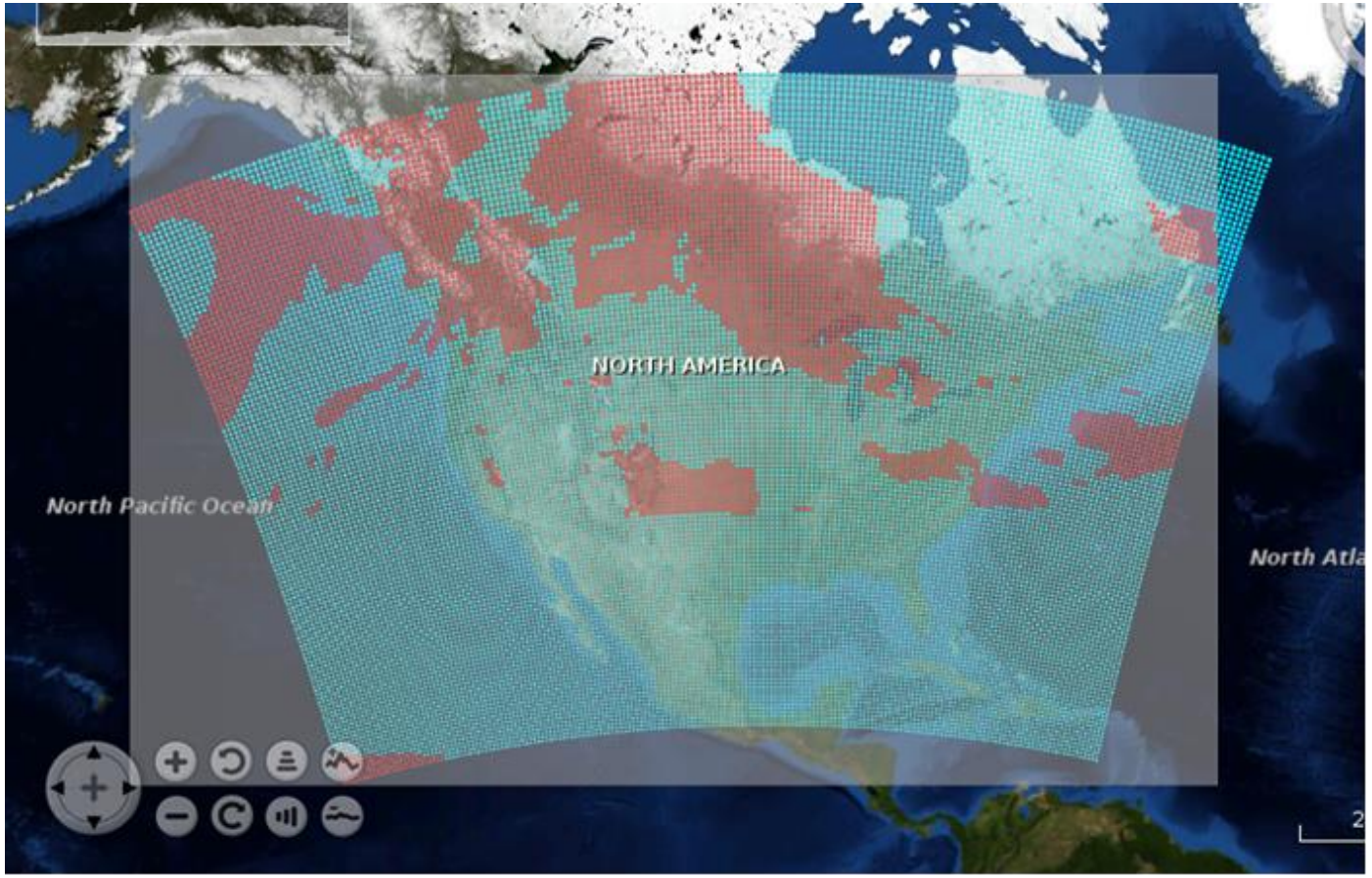
3 Performance Evaluation

4

5

Using Thunderstorm Indices

- Best 4 layer Lifted Index
- Storm Relative Helicity
- Convective Inhibition



23866 Input Data Records
Execution Time : 15.75 seconds

ML Based Clustering

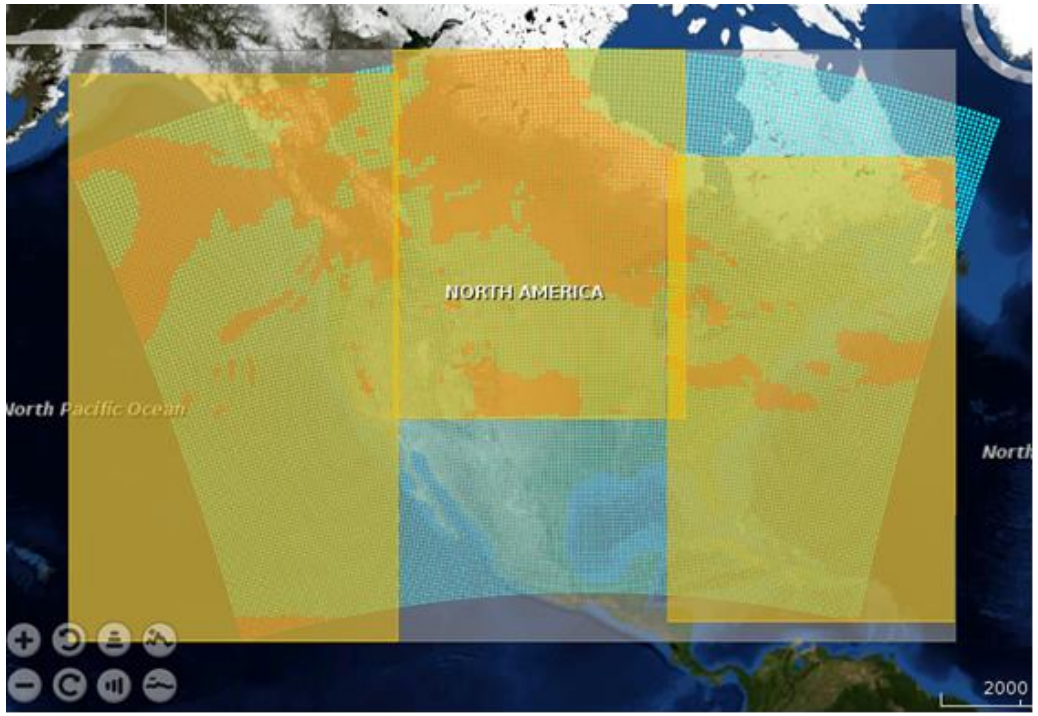
LARIDAE
ENHANCED WEATHER MONITORING SYSTEM

3 Performance Evaluation

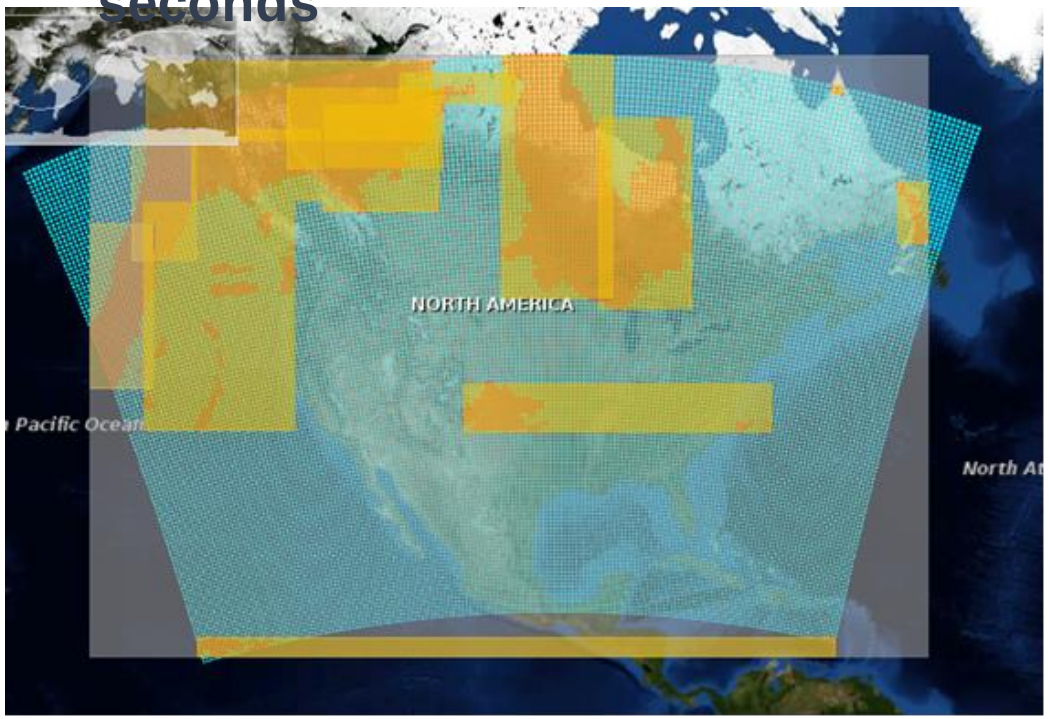
4

5

K-Means Algorithm : 19.299 seconds



GMM Algorithm : 63.063 seconds



Results Comparison

LARIDAE
ENHANCED WEATHER MONITORING SYSTEM

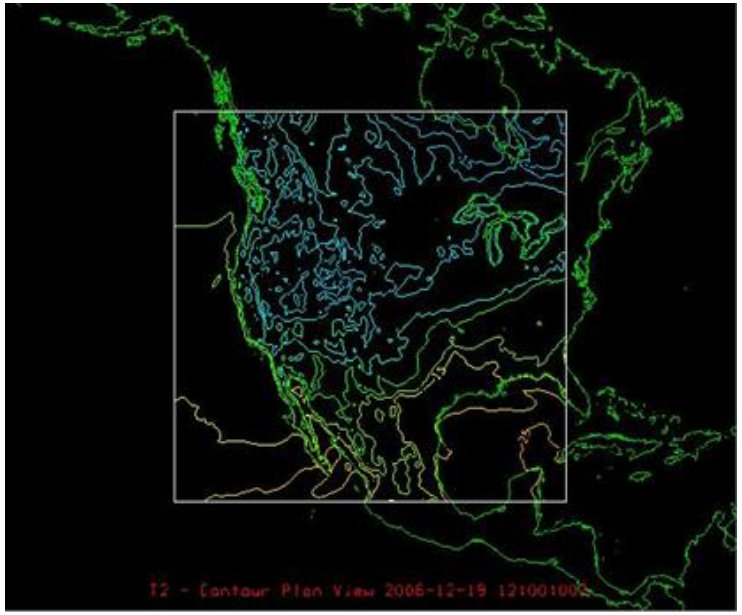
3 Performance Evaluation

4

5

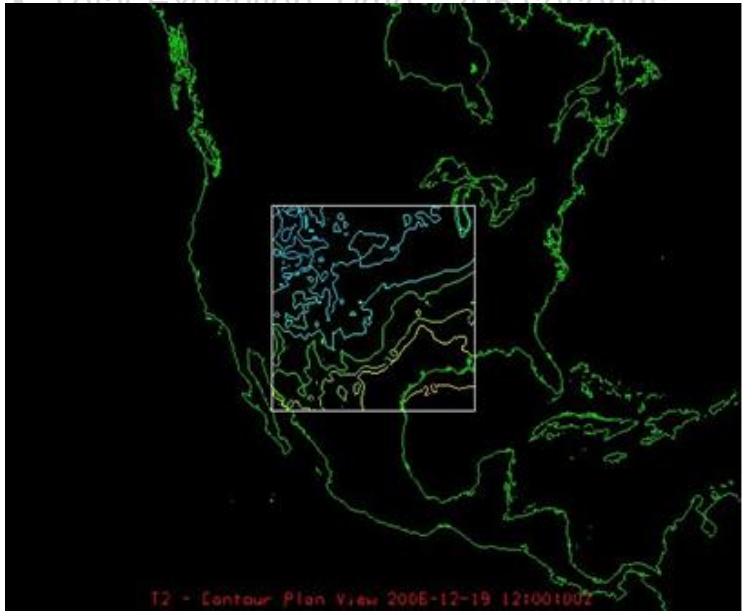
Scenario 1

- Without Preprocessing
- Total Execution Time :745 seconds



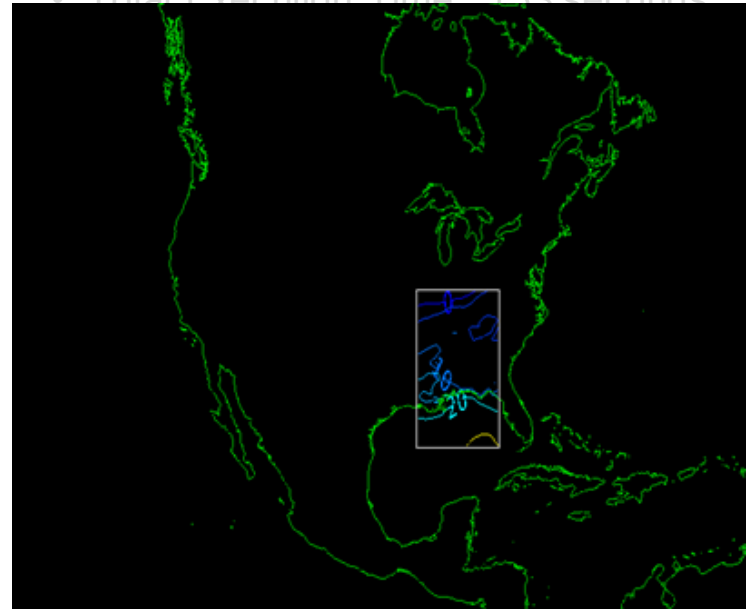
Scenario 2

- Clustering using K Means
- WRF Execution Time : 165.632 seconds
- Total Execution Time : 305 seconds



Scenario 3

- Clustering using GMM
- WRF Execution Time : 73.978 seconds
- Total Execution Time : 153 seconds



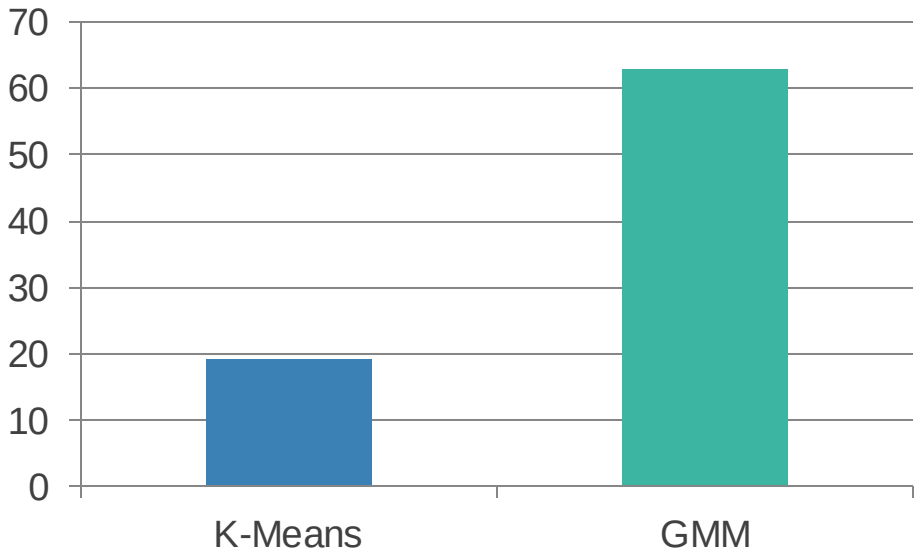
Execution Time Comparison

LARIDAE
ENHANCED WEATHER MONITORING SYSTEM

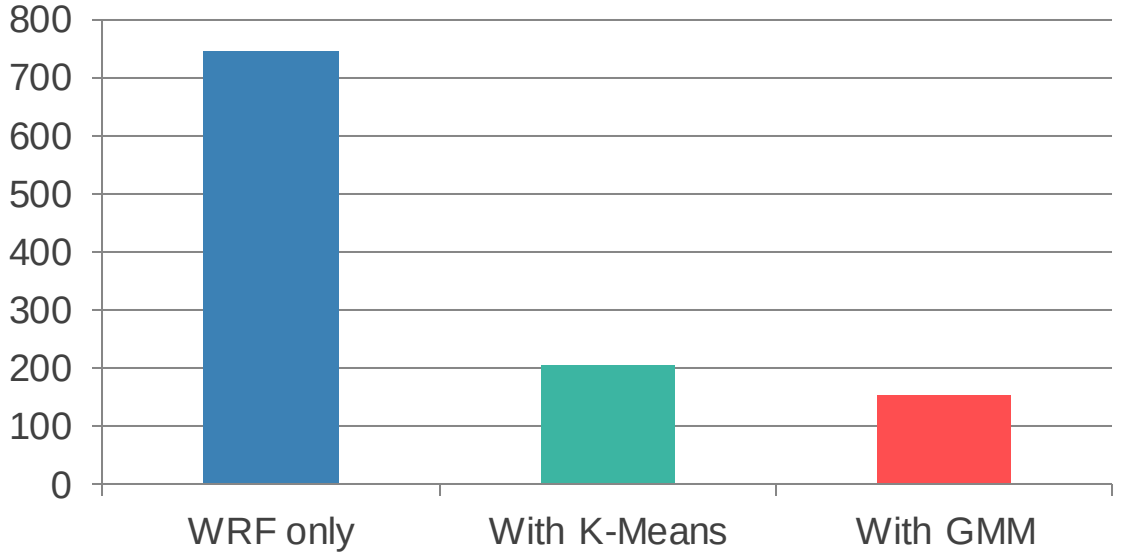
3 Performance Evaluation

4

5



ML Algorithm Execution time (Seconds)



Total Execution Time (Seconds)

- Significant Time Gain
- GMM : Expensive Algorithm But Gives Better Performance in Clustering

FUTURE IMPROVEMENTS

Enhancements

LARIDAE
ENHANCED WEATHER MONITORING SYSTEM

5 Future Improvements

- Proof of concept
- Identifies only the thunderstorm signatures
- Other weather indices should be added
- Better algorithm to find “k” value in GMM clustering algorithm
- Feeding streams of data from different weather data sources
- Adapting the input data collection resolution dynamically wrt the feedback from individual components

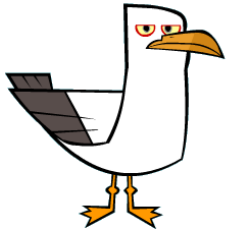
Conclusion

LARIDAE
ENHANCED WEATHER MONITORING SYSTEM

5 Future Improvements

- We have built a successful prototype of the solution concept
- Statics shows a considerable improvement of the running time of WRF module when integrated with our system
- Total running time of the forecasting procedure has reduced by 75%-85% without affecting the accuracy of the results
- The solution is feasible, cost-effective and can be further extended

THANK YOU!



chamil.11@cse.mrt.ac.lk

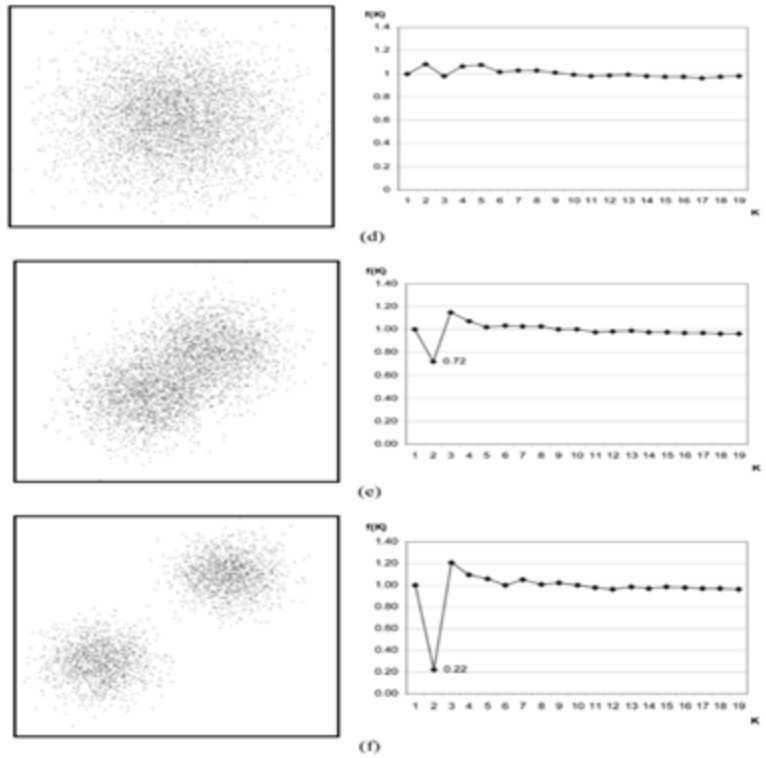
BACKUP SLIDES

ML Clustering Layer

Standard K-Means for Clusterin

- Faster compared to other clustering algorithms
- Need to specify ‘K’

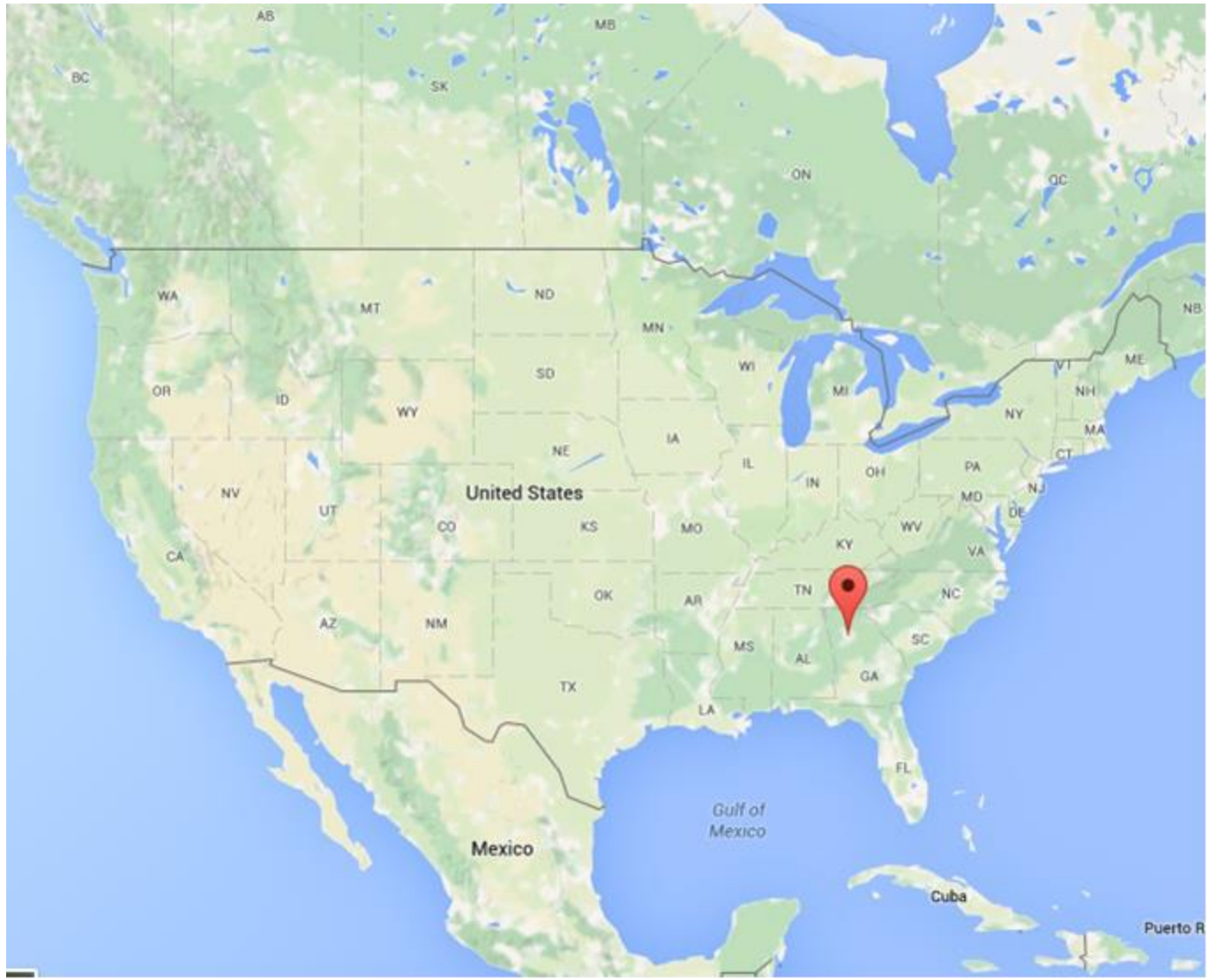
Finding the ‘K’

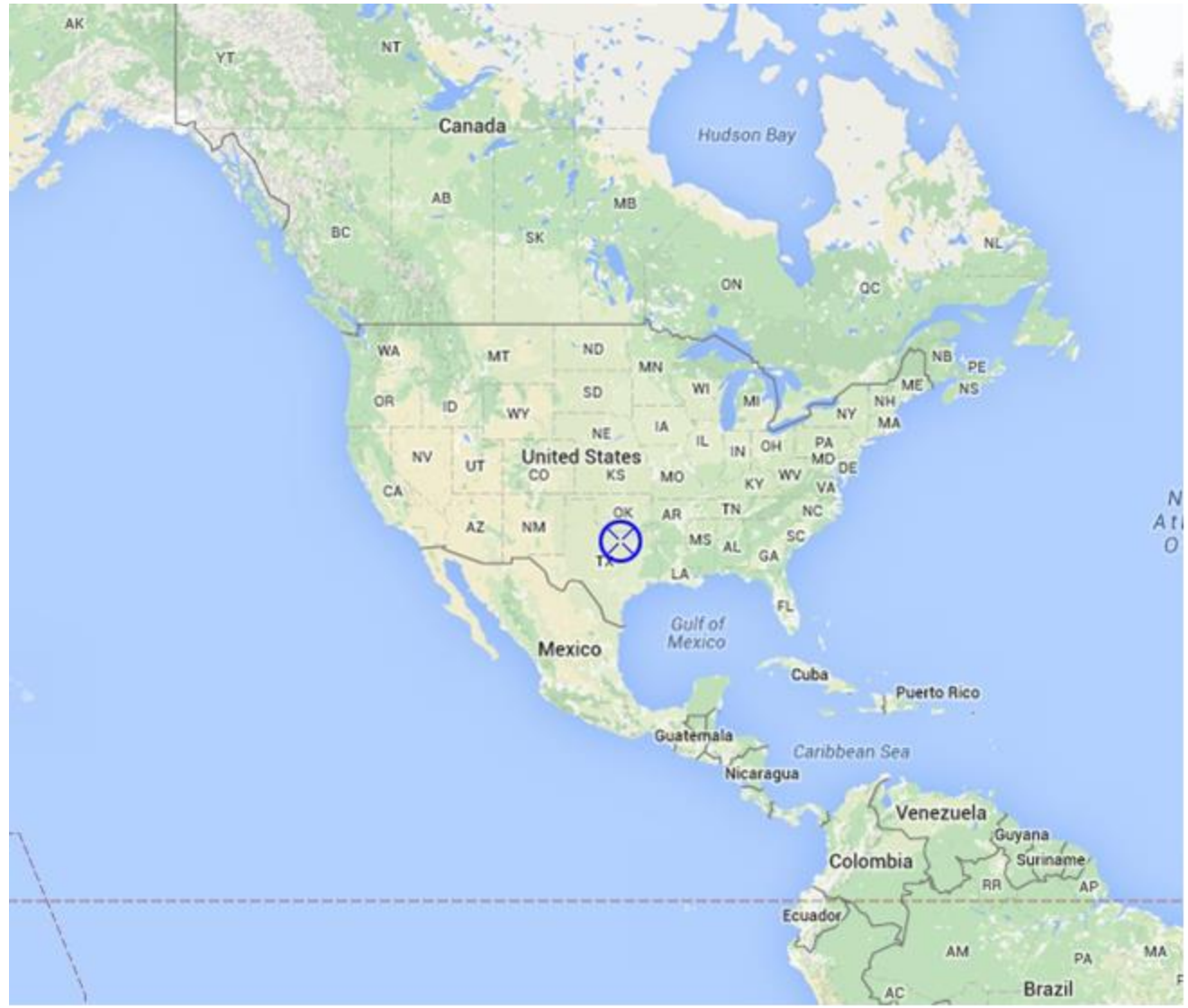


Solution :
Improved K-Means

Event Details

Event	**Thunderstorm Wind**
Magnitude	**50 kts.**
State	**TEXAS**
County/Area	**WISE**
WFO	**FWD**
Report Source	**Department of Highways**
NCDC Data Source	**CSV**
Begin Date	**2006-12-20 06:00:00.0 CST-6**
Begin Location	**9SSW DECATUR**
Begin Lat/Lon	**33.11/-97.54**
End Date	**2006-12-20 06:00:00.0 CST-6**
End Location	**9SSW DECATUR**
End Lat/Lon	**33.11/-97.54**
Deaths Direct/Indirect	**0/0** (fatality details below, when available...)
Injuries Direct/Indirect	**0/0**
Property Damage	**10.00K**
Crop Damage	**0.00K**
Episode Narrative	**An upper level low pressure system and associated Pacific front moved through North Texas in the early morning hours.**
Event Narrative	**A few signs were blown down along Highways 114 and 151 near the leading edge of a shallow line of thunderstorms.**


AB
MB
BC
SK
ON
QC
WA
MT
ND
MN
NB
WI
VT
ME
OR
ID
SD
MI
NH
WY
NY
MA
IA
NE
IL
CT
PA
OH
NV
United States
IN
NJ
UT
MD
DE
CO
KS
MO
WV
KY
VA
CA
TN
NC
OK
AR
AZ
NM
MS
SC
AL
GA
TX
LA
FL
Gulf of
Mexico
Mexico
Cuba
Puerto R

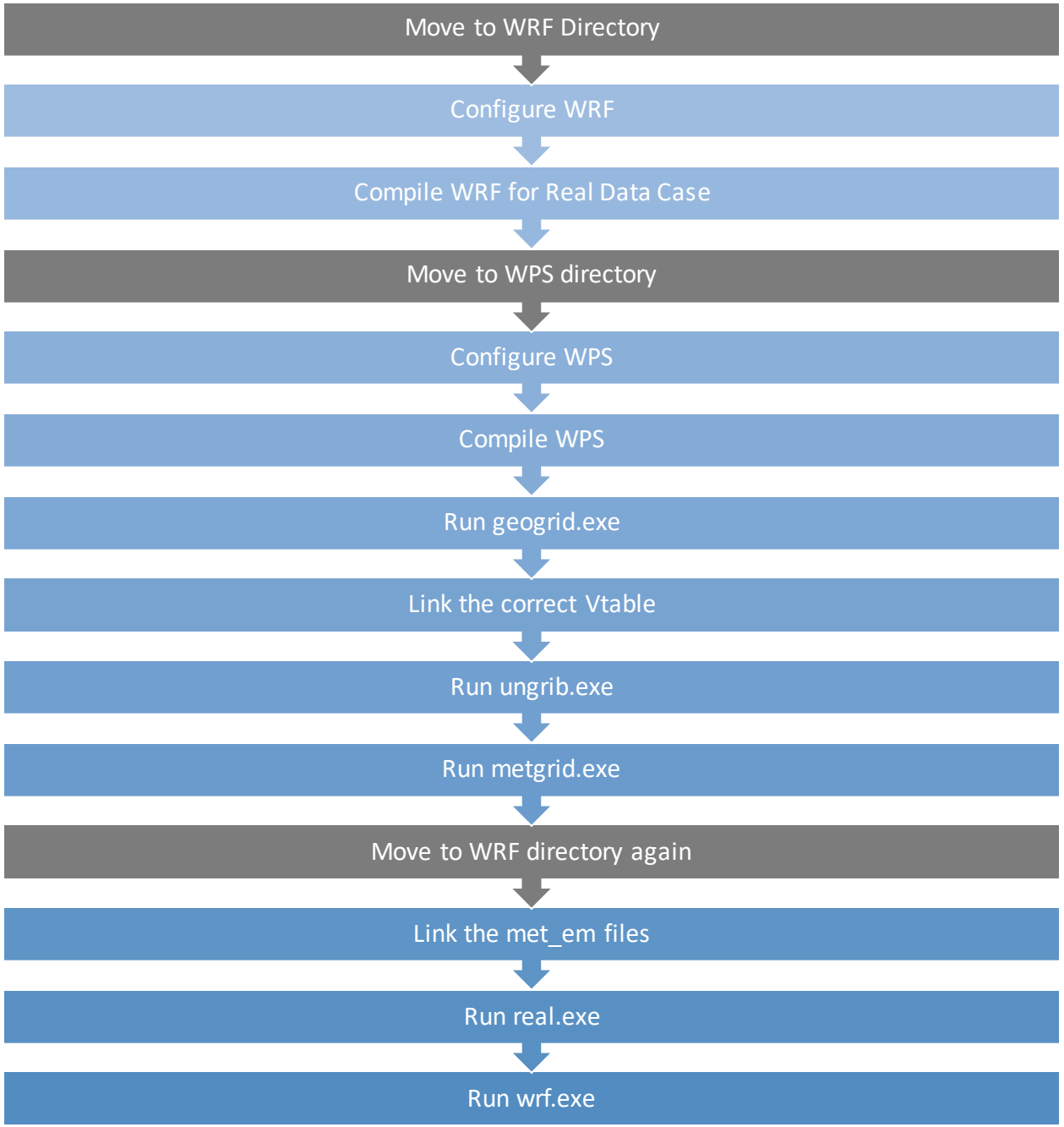

AK
YT
NT
Canada
Hudson Bay
AB
MB
BC
SK
NL
ON
QC
WA
MT
ND
MN
NB
PE
WI
ME
NS
OR
ID
SD
MI
NH
WY
NY
MA
IA
NE
IL
IN
OH
PA
NV
UT
United States
CO
KS
MO
KY
WV
MD
DE
CA
VA
OK
AR
TN
NC
AZ
NM
MS
AL
SC
GA
TX
LA
FL
Gulf of Mexico
Mexico
Cuba
Puerto Rico
Guatemala
Caribbean Sea
Nicaragua
Venezuela
Guyana
Colombia
Suriname
RR
AP
Ecuador
AM
PA
MA
AC
Brazil

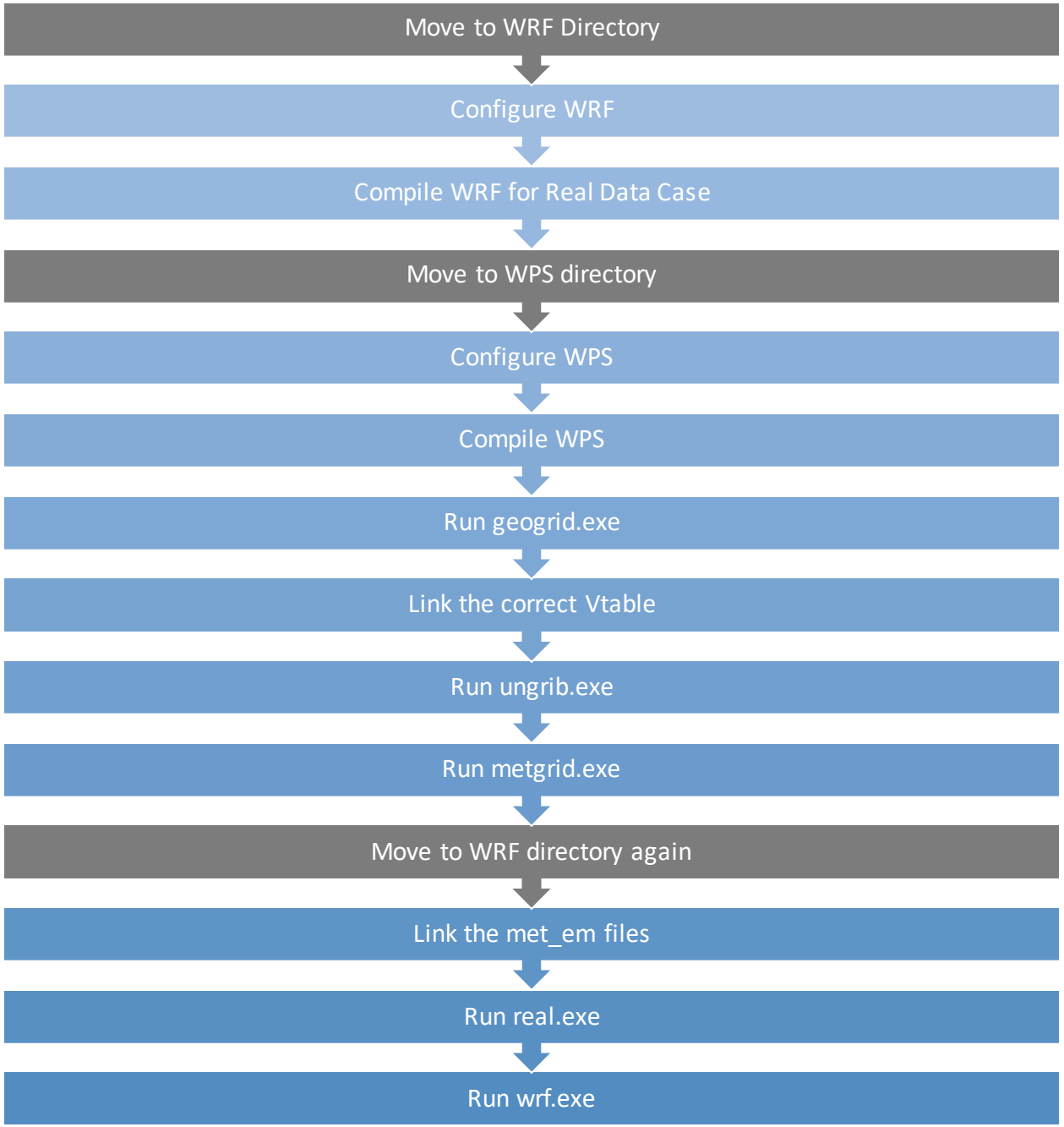
```
&share
wrf_core = 'ARW',
max_dom = 2,
start_date = '2008-03-24_12:00:00','2008-03-24_12:00:00',
end_date = '2008-03-24_18:00:00','2008-03-24_12:00:00',
interval_seconds = 21600,
io_form_geogrid = 2
/

 &geogrid
parent_id = 1, 1,
parent_grid_ratio = 1, 3,
i_parent_start = 1, 31,
```

Figure 4.18 - namelist.wps file contents

```
j_parent_start = 1, 17,
s_we = 1, 1,
e_we = 74, 112,
s_sn = 1, 1,
e_sn = 61, 97,
geog_data_res = '10m','2m',
dx = 30000,
dy = 30000,
map_proj = 'lambert',
ref_lat = 34.83,
ref_lon = -81.03,
truelat1 = 30.0,
truelat2 = 60.0,
stand_lon = -98.,
geog_data_path = '/mmm/users/wrfhelp/WPS_GEOG/'
/
&ungrib
out_format = 'WPS',
prefix = 'FILE',
/
&metgrid
fg_name = 'FILE'
io_form_metgrid = 2,
/
```


Move to WRF Directory
Configure WRF
Compile WRF for Real Data Case
Move to WPS directory
Configure WPS
Compile WPS
Run geogrid.exe
Link the correct Vtable
Run ungrib.exe
Run metgrid.exe
Move to WRF directory again
Link the met_em files
Run real.exe
Run wrf.exe


Move to WRF Directory
Configure WRF
Compile WRF for Real Data Case
Move to WPS directory
Configure WPS
Compile WPS
Run geogrid.exe
Link the correct Vtable
Run ungrib.exe
Run metgrid.exe
Move to WRF directory again
Link the met_em files
Run real.exe
Run wrf.exe

What we achieved

LARIDAE
ENHANCED WEATHER MONITORING SYSTEM

5 Future Improvements